stuck. Students often spend a lot of time (sometimes more than one hour) searching for the cause of the problem, or a way to overcome the problem, before they decided to ask the teacher for help. Two student pairs did not complete the project at all. On the basis of this experience, it was concluded that lagging behind should be avoided by all means. Also, it was decided that it is absolutely necessary for students to do the geodata visualization and processing in GIS at school, under the supervision of a teacher. Next, it was decided that it is necessary to include strict deadlines in the planning of the project for handing in inquiry plans and collecting geodata. Problems due to lagging behind of students also occur in traditional geographic inquiry projects. However, in the project ‘Services and Customers’, such problems seemed to be more pronounced, as the success of each phase depends on the success of the previous one.

11.2.3.2 Collected data

Five types of data were collected during and shortly after the tests. First, the inquiry-planning phase of Test II was videotaped, with a focus on instances in which students were formulating inquiry plans together, and instances in which the students and teacher discussed the inquiry plans. The presentations in Test II and the whole-class discussion in Test II were also videotaped. Second, field notes were made during the GIS training session and the GIS phase. All problems were registered. Third, on-task mini interviews were conducted with students in Tests II and when they were working with GIS. Fourth, all inquiry plans, presentations, and reports constructed by students were collected. Fifth and finally, surveys were conducted with the students immediately after they received their grades in each of the tests.

11.2.4 Evaluation Stage II

11.2.4.1 Set-up and learning goals of the project

Although it was clear that the project was not yet viable and effective, the research team was satisfied with the learning goals of the second design. The same applied to the set-up. It included all necessary phases and had a logical structure. It was expected that the learning goals could be reached by making revisions in the design of the individual phases.

11.2.4.2 Evaluation of the inquiry-planning phase

Students’ initial inquiry plans and the videos and on-task mini interviews of the inquiry-planning phase showed that many students found it difficult to construct inquiry plans with a high domain-specific quality. Six types of problems as a result of inquiry strategy planning-related learning difficulties were noticed by the research team. Now follows a short description of these problems.

First, students often said they did not know what to do and asked for more guidance. Some students were not able to construct research plans by themselves. So apparently, students found it difficult to understand what they were supposed to do and how they were supposed to do it. They did not completely understand the overall inquiry procedure of the project. On the basis of this experience, it was concluded that showing the intended end products in the introduction phase and practicing the relevant GIS operations (in another context) in the GIS
training session may have helped to avoid problems caused by this learning difficulty, but that this is not enough. More is needed in order to ensure that students construct good inquiry plans by themselves\textsuperscript{PAC15}.

Second, students’ inquiry plans often did not contain lists of inquiry questions, hypotheses, and survey questions, but contained textual descriptions instead\textsuperscript{BP5} (see, for example Table 11-6). About one quarter of the inquiry plans lacked any survey questions. So, apparently, students often find it difficult to construct clear and complete inquiry plans\textsuperscript{LD22}. It was therefore concluded that it is necessary to give students a sheet with the requirements of the inquiry plans\textsuperscript{DP19}.

\textbf{Table 11-6: Fragment from an inquiry plan constructed by students in Test II}\textsuperscript{A}

<table>
<thead>
<tr>
<th>Text in the student inquiry plan</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>We are going to ask for the postal codes in order to map where the customers live. [...] As we want to know why they visit this shop we are also going to ask them for the reason (such as buying something special, or just looking around). The accessibility may also affect the market area. [...] Therefore we want to know how people come to the city centre and we are going to ask for their transport means.</td>
<td>The inquiry plan does not contain a clear list of survey questions.</td>
</tr>
</tbody>
</table>

Third, many student inquiry plans contained selections of services, inquiry questions, and hypotheses that did not connect to each other (see, for example, Table 11-7). Also, they often contained survey questions that were ineffective for testing the hypotheses\textsuperscript{BP6} (see, for example, Table 11-7). So apparently, students find it difficult to construct inquiry plans with a high external validity\textsuperscript{LD23}. The videodata from Test II\textsuperscript{C} showed that students often found it difficult to formulate hypotheses in general (see, for example, Table 11-8).
Chapter 11: Thick description

Table 11-7: Simplified version of an inquiry plan constructed by two students who investigated the market areas of three Bakker Bart bakeries in Nijmegen (Test II(B))

Main inquiry question

➔ Why are there differences in the size of the market areas of three Bakker Bart bakeries?

Sub-questions

1. Where do the customers of the three bakeries come from?
2. What is the difference in the location phenomena at the three bakeries?
3. What is the difference in the amount of advertising of the three bakeries?
4. What are the differences in the assortment?
5. Are there differences in the customers because of the differences in the assortment?

Hypothesis

➔ The size of the market area is, from large to small: (1) Bakker Bart Nijmegen, (2) Bakker Bart Beuningen and (3) Bakker Bart Groesbeek.

Data collection strategy (survey questions)

Survey question 1: What is the aim of your visit?
   A) I got hungry when I was shopping
   B) I’m just having a look
   C) I travelled to this bakery to buy bread
   D) For lunch
   E) Other

Survey question 2: What do you value most about Bakker Bart?
   A) The sandwiches
   B) The service
   C) The price-quality ratio
   D) The large assortment
   E) Other

Survey question 3: What is your opinion on the special offers of the bakery (on a 1 to 10 scale)?

Survey question 4: What is your postal code?

Table 11-8: Transcript from a one-on-one discussion in the inquiry-planning phase in Test II(C) (A)

<table>
<thead>
<tr>
<th>Line</th>
<th>Actor</th>
<th>Text / Action</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>“What should be in the hypothesis? I still do not get it. [...] What do we have to do?”</td>
<td>S does not know the meaning of the concept ‘hypothesis’</td>
</tr>
<tr>
<td>2</td>
<td>TC</td>
<td>“It is a statement about, well, then you say what you expect to find. You are going to investigate supermarkets right, so what do you expect? Why did you choose those supermarkets by the way?”</td>
<td>TC explains what is a hypothesis, and tries to stimulate students to formulate implicit hypotheses</td>
</tr>
<tr>
<td>3</td>
<td>S</td>
<td>“Well, the Aldi is very different from the Albert Heijn. The customers are different, more like poor people. [...] We want to go to the biological supermarket too, because I think that it attracts people from far away. There are two biological supermarkets in Zutphen, but none in the villages”</td>
<td>S formulates two hypotheses without being aware of it: one about the differences in per capita income of the customers at three services; and one about the size of the market areas of those three services.</td>
</tr>
<tr>
<td>4</td>
<td>TC</td>
<td>“That’s a hypothesis! Just like you say: The biological supermarket attracts people from far away. So you expect that people have travelled much further. A hypothesis. You could write that down in your inquiry plan”</td>
<td>TC makes one of these hypotheses explicit.</td>
</tr>
</tbody>
</table>

Notes: TC = Teacher C; S = Student.
Fourth, students’ inquiry plans often had a weak spatial perspective. Four student pairs wanted to compare services that had similar values for the spatial properties. Next, more than half of the students formulated inquiry questions and hypotheses that related non-spatial phenomena to each other. For instance, one student pair who wanted to investigate music stores formulated hypotheses about the relationship between the age (of customers) and the music preferences (of those customers), and between the music preferences (of customers) and the frequency that (those customers) download music from the Internet. The students formulated survey questions such as “What kind of music do you like?”, “What is your age?”, and “How often do you download music?” Although this could be an interesting inquiry, it could not be an interesting geographic inquiry. Table 11-7 presents an inquiry plan of two students who investigated three bakeries of the Bakker Bart branch. Although this is an interesting inquiry plan, it also has some flaws. For instance, the second survey question asked customers what they value most about Bakker Bart Bakeries in general. This survey question provides little insight into why people chose to visit this specific bakery and not another bakery. Therefore, it offered little insight into the causes of the differences in the size of the market areas. Six student pairs did not even mention the size of market areas in their initial inquiry plan. In summary, it can be concluded that the inquiry plans were not very geographic because they contained few spatial key concepts (for the definition of ‘spatial key concept’, see Section 5.4.3). So, apparently, students are often not aware that the location matters when they select services, formulate inquiry questions and hypotheses, and construct surveys. Table 11-9 shows a transcript from a one-on-one (student and teacher) discussion in Test II that shows that students did not see the use of following a spatial perspective when selecting services. As the property ‘size of the market area’ is connected to a spatial key concept, including this property in the inquiry questions or hypotheses is the first key to making students’ inquiry plans more spatial, and thus more geographic.

Table 11-9: Transcript from a one-on-one discussion in the inquiry-planning phase in Test IIc (B)

<table>
<thead>
<tr>
<th>Line</th>
<th>Actor</th>
<th>Text / Action</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TC</td>
<td>“So you want to compare a Kiosk [news paper stand], Burger King, and Swirls [ice cream stand] at the railway station in Arnhem. Those are very different categories of services at the same place. I suggest you, well... Why do you want to visit those services? And why Arnhem?”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>“That will be fun! [...] And it is very convenient, because my mother is the manager of the Swirls [the ice cream stand], so we could ask her if she could conduct the surveys for us”</td>
<td>S did not see the use of following a spatial perspective in their inquiry</td>
</tr>
</tbody>
</table>

Notes: TC = Teacher C; S = Student.

Fifth, students’ inquiry plans often did not contain the right type of inquiry questions. Although the teacher advised students not to formulate open survey questions that ask for text, some of students still did so. Some other students discovered during the geodata collection that people could give multiple answers at the same time. This learning difficulty connects to the internal validity of the surveys. So apparently, students often find it difficult to formulate the right type of survey questions, and to construct surveys with a high internal validity.
Sixth and finally, students’ inquiry plans often contained incorrect survey questions. For instance, the multiple choice answers which accompany the survey question “What do you value most about this bakery?” in the inquiry plan about the Bakker Bart bakeries (Table 11-7) are not mutually exclusive. People could buy sandwiches because the bakery has a good price/quality ratio. So apparently, students often found it difficult to formulate the correct survey questions. This learning difficulty also connects to the internal validity of the surveys.

The research team concluded that the domain-specific quality of the inquiry plans of students in Test Stage II was far too low and that students need more support. Not only should the teacher check the inquiry plans before students start collecting geodata, students should construct inquiry plans in the lesson, under supervision of a teacher who could support them so they are able to overcome the problems as a result of the inquiry strategy planning-related learning difficulties mentioned above.

Besides problems caused by the learning difficulties mentioned above, some unexpected good things happened too. For example, one student pair in Test II told the teacher that they wanted to collect geodata about the customer flow at their services (a cheap supermarket, an expensive supermarket, and an organic supermarket). They wanted to count the number of customers in a 15-minute time span, and asked the teacher if they could visualize such geodata in GIS maps. This is possible with GIS, but the teacher did not know how to do so. He subsequently asked the researcher for help, who explained to the students how they could visualize such geodata.

**11.2.4.3 Evaluation of the Excel phase**

The field notes provided some insight into the nature of the problems that occurred when students entered, manipulated, and aggregated their geodata in Excel. Most problems were related to deficiencies in the design of the Excel handout. The handout contained about ten errors, obscurities, and omissions. For example, the worked-out example told students to enter the formula ‘=average(customers!H2:H21)’. Many students copied the formula including the apostrophes, and subsequently got stuck. Besides this, the Excel part of the handout also contained some inappropriately designed tasks. For example, in the next step, the handout told students to change the references in the formula for the average. Many students did not know how to do that, and asked the teacher for help. So apparently, asking students to copy a formula and change the references is an inappropriately designed task, and it is probably more viable to ask students to enter the formula ‘=average(’, and then to click the right cells.

**11.2.4.4 Evaluation of the GIS phase**

Four types of problems occurred in the GIS phase. The first type of problems were caused by structural imperfections in students’ Excel files. About 90 per cent of the students got stuck at least once because their geodata tables did not meet the criteria of a proper database table. It was discovered that students had not entered, manipulated, and/or aggregated their geodata correctly in the Excel phase, because of deficiencies in the Excel handout, and/or because the students had not read the worked-out examples in the Excel handout carefully. So apparently, students found it difficult to read the worked-out examples in the Excel handout precisely, and found it difficult to understand the importance of constructing tables which meet the criteria of a proper database table. Unfortunately, the
GIS software did not give clear error messages. Instead, GIS tools simply did not work or produced erroneous results. Students therefore had no idea that these problems were caused by structural imperfections in their geodata file.

The second type of \( \alpha \) problems were caused by content-related imperfections in students’ Excel files. Around 20 per cent of the students discovered during the GIS phase that their maps were “strange” or “incorrect”: customers were displayed somewhere in France; market areas covered the entire continent; or pie chart maps displayed unrealistic values. It was discovered that these students had not entered the formulas correctly in their files in the Excel phase\(^{\text{P9}}\). Again, this was caused by imperfections in the Excel handout, and/or caused by the learning difficulty that students often find it difficult to read worked-out examples in the Excel handout precisely\(^{\text{P10}}\). Besides this, about 30 per cent of the students discovered during the GIS phase that they could not visualize all their survey geodata in GIS maps, because their Excel files were incomplete\(^{\text{P9}}\). It was discovered that these students had not entered, manipulated, or aggregated all their geodata in Excel. In the tests, it was noticed that students found it difficult to follow the Excel handout right to the end, and wanted to start visualizing their geodata in GIS as soon as possible\(^{\text{P12}}\). About half of the students who encountered problems as a result of content-related imperfections in their Excel files asked the teacher and researcher for help. The other students were able to identify the cause of the problems themselves. When these students tried to fix the problem by opening and editing their geodata tables in Excel, they discovered that their map layers became corrupted\(^{\text{P13}}\). This is because these map layers are temporary map layers based on a table. Apparently, students found it difficult to understand the nature of temporary map layers, and found it difficult to understand that it is risky to work with geodata tables in Excel and GIS at the same time\(^{\text{P13}}\).

Some of the \( \alpha \) problems mentioned above could be avoided by revising the GIS software. First, educational GIS software should have less stringent requirements for a proper database table\(^{\text{P21}}\), or give clear error messages and solutions when students try to add geodata tables that do not meet the criteria of a proper database table\(^{\text{P22}}\). Second, educational GIS software should create permanent map layers instead of temporary map layers when students visualize geodata tables with XY geodata, in order to avoid the problem that students start working with the same data in Excel and GIS at the same time\(^{\text{P23}}\). As software development was not part of the design process, it was concluded that the design of the GIS handouts and the Excel handouts should take the imperfections in the GIS software into account, and anticipate future \( \alpha \) problems\(^{\text{BC1}}\). This means that it is absolutely necessary to give students an Excel handout with a high degree of guidance that consists of step-by-step instructions with worked-out examples to guide them into constructing geodata tables which meet the criteria of a proper database table\(^{\text{P24}}\).

The third type of \( \alpha \) problems were related to the design of the GIS handout. There were four causes for these problems. First, the GIS handout did not have an appropriate structure. Few students actually followed the step-by-step instructions in the handout right to the end. Most students skipped steps and then got stuck. Besides this, students often did not use the GIS training handout as a reference book, as planned when they did not know how to perform a GIS operation. Instead, they just asked the teacher for support. On the basis of this experience, the research team concluded that a semi-guided design in which students follow step-by-step instructions with references to the GIS training handout is totally unsuitable\(^{\text{BC14}}\). Second, the design of the tasks was at several points inappropriate. For example, the GIS handout instructed
students to construct layouts in GIS, but the layout menu turned out to be far too complex, and
many students got stuck. So apparently, letting students construct layouts in GIS was an
inappropriately task, and it is probably more viable to instruct students to make screen dumps
of the GIS environment and to edit these screen dumps in PowerPoint. Third, the order of the tasks
was at several points inappropriate. For example, the GIS handout instructed students to first set
the labels for the map layer ‘services’, then copy the map layer several times, and, finally, edit
the symbology of the map layers so that each one shows a different attribute. This resulted in 5
to 8 stacked map layers with the same labels. Many students asked how to get rid of the
duplicate labels. They did not unset the labels for each map layer themselves. So apparently,
the order of the tasks was inappropriate, and it is probably more viable to instruct students to
first copy the map layer ‘services’ several times, then edit the symbology of the map layers, and,
finally, set the labelling for the original map layer. Fourth and finally, it was discovered that the
GIS handout and GIS training handout contained tens of tiny errors, obscurities, and omissions.
Although many of them seemed innocent, they all resulted in students getting stuck. For
example, many students got stuck when they followed the instruction in Figure 11-5. This
handout instructed them to copy one of the pie chart map layers and to change the symbology
and create a quantities map layer. Some students changed the symbology of the old pie chart
map layer instead of the new pie chart map layer. As the old one was ‘below’ the new one,
they could not see the results of the operation, and thought something was wrong. So the
handout should have instructed students to change the symbology of the new map layer.

The fourth and final type of a problems consisted of problems which also occurred in Test
Stage I: about 20 per cent of the students had saved their map layers locally\textsuperscript{93}, and about 10
per cent of the students had moved or renamed their map layers with Windows Explorer\textsuperscript{94},
although the teacher and researcher had clearly told them not to do so. It was therefore decided
to provide students a GDB file, and to revise the instruction for saving map layers. The revised
handout would tell students to save map layers as feature class files (which can only be saved
in a GDB file) instead of shapefiles (which can be saved everywhere).

The two handouts were revised, tested, and evaluated again in Cycle III and IV. In the course of
the design process, it was learned that the Excel handout and the GIS handout should be
absolutely perfect in order to ensure that students are able to visualize and process all their
geodata without getting stuck. The handouts should have: (1) a suitable structure; (2)
appropriately designed tasks; (3) appropriately ordered tasks; and (4) instructions which are 100
per cent free of errors, obscurities, and omissions\textsuperscript{95}. Regarding the structure, it was learned
that step-by-step instructions with worked-out examples were best for the project ‘Services and
Customers’. The revision, testing, and evaluation of the handouts also resulted in tens of mini-
design principles for an appropriately designed tasks and appropriately ordered tasks. Because
this is not the main interest of this dissertation, this part of the design process, and the mini-
design principles that resulted from it, is not reported in this dissertation.

As well as the a problems, the research team also noticed a b problem: none of the maps in the
student representations and student reports contained two map layers that showed a
relationship\textsuperscript{96}. Instead, the presentations and reports contained a series of maps in which each
map showed one map layer only, often in the chronological order in which they were
constructed (see, for example, Figure 11-6). So, apparently, students find it difficult to
understand that constructing maps is part of a process that is directed towards answering
inquiry questions and testing hypotheses, and not an aim in itself\textsuperscript{19}. It was therefore
concluded that teachers stimulate students to include maps in their presentations and reports that show the answer to their inquiry question or show whether their hypothesis should be accepted or rejected\[^{DP25}\].

Figure 11-6: Simplified versions of maps constructed by two students who investigated the market areas of three Bakker Bart bakeries in Nijmegen (Test II\(^{B}\))
A major threat to the viability of the inquiry project was that the teachers were able to help only around 50 per cent of the students who encountered α problems. The researcher helped out the other 50 per cent. In the evaluative talks, the three teachers said that they were able to follow the Excel and GIS handouts and enter, manipulate, aggregate, and visualize the geodata themselves, but that they did not have enough knowledge to diagnose all types of α problems, and that they did not have enough knowledge to take over the task and overcome the problems, or to support students so that they are able to overcome the problems. This especially applied to those α problems caused by structural imperfections in students’ Excel files. This is no surprise, as the teachers had not been trained in supporting students when they work with GIS. However, during the tests, the teachers had shown that they were willing to learn how to support the students, so that they could do better next time. They had frequently asked the researcher to show the cause of the problems, and to explain how these problems can be overcome. Still, the teachers argued that the number of α problems should reduce considerably in order to increase the viability of the project.

### 11.2.4.5 Evaluation of the presentation phase

Most of the presentations were easy to follow. However, there were a few drawbacks. First, the presentations took about twice as long as planned. As a result, there was little time for one-on-one (teacher and student) discussions. Second, many students described the procedures in detail, and the teachers and the other students found it boring to hear the same stuff over and over again. Besides these drawbacks, the research team also noticed that most presentations contained a reflection section, but that almost all utterances of reflection consisted of opinions on the inquiry project (the degree to which students found the project interesting, useful, and/or fun), the collaboration between the students (whether collaboration went fine or not), and the degree of difficulty. Only two presentations contained a note of reflection on the domain-specific quality of their inquiry project. So apparently, students found it difficult to reflect on the domain-specific quality of their inquiry project by themselves.

After the presentations, the teachers tried to raise students’ geographic thinking to a higher level. Teacher A, B, and C applied, respectively, 5.7, 2.5, and 2.9 interventions per student pair. Most teacher interventions consisted of questions that asked the students to explain their geodata further (declarative knowledge in Domain Da). Teacher C also asked many questions which focused on superficial details of the layout of the maps (procedural and strategic knowledge in Domain C). Hardly any teacher interventions focused on the inquiry questions, hypotheses, and geodata collection strategy (procedural and strategic knowledge in Domain A and Bb). Only one student pair had used the theoretical insights from the literature in their presentations, and none of the teacher interventions aimed to link these theoretical insights to students’ geodata.

In the evaluative talks, the teachers said that students’ presentations offered many opportunities for discussions on the subject of services and customers, and to let students reflect on the domain-specific quality of their inquiry project. However, they also said that they found it difficult to ask sensible questions about students’ inquiry strategy. Some inquiry projects were so bad (incomplete and unclear, weak spatial perspective, low internal and external validity, etc.) that teachers did not know where to start. They therefore primarily focused on the content. Table 11-10 presents a transcript from a discussion between the teacher and the two students.
who investigated the Bakker Bart bakeries. The transcript shows how teacher B1 supported students to identify a relationship. However, it is questionable whether the teacher succeeded in doing so. The transcript was interpreted in Cycle V with the help of the domain-specific construct for use in educational settings (see Section 10.3), which was not completed at the time of the Cycle II.

Table 11-10: Transcript from a one-on-one discussion at the end of a presentation of two students who investigated the market areas of three Bakker Bart bakeries in Nijmegen (Test II)

<table>
<thead>
<tr>
<th>Line</th>
<th>Actor</th>
<th>Text / Action</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TB1</td>
<td>“I have a question about what you have just said: that there are more people from the villages who visit the Bakker Bart Bakery in Groesbeek. You have also said something about other bakeries and supermarkets. What did you mean?”</td>
<td>TB1 aims to let students find out that the size of market areas of bakeries is related to the distance to competitive bakeries.</td>
</tr>
<tr>
<td>2</td>
<td>S1</td>
<td>“Well, it is because… In Beuningen there are few people from the villages. Maybe they go to the supermarket more often to buy bread, and not to the Bakker Bart Bakery”</td>
<td>S1 hypothesized that there is a difference between Beuningen and Groesbeek regarding the ratio between the amounts of bread bought in bakeries and the amount of bread bought in supermarkets. Although this could be true, it does not explain the differences in the size of market areas.</td>
</tr>
<tr>
<td>3</td>
<td>TB1</td>
<td>“OK. But what do you have to do to find out whether that is really the case?”</td>
<td>TB1 asked S1 how he could check this hypothesis.</td>
</tr>
<tr>
<td>4</td>
<td>S1</td>
<td>“Uhmmm… You should have a look at the transport means that people use. Old people cannot go by car to the bakery. The distance is too far. So maybe they go to another bakery that is nearer. And use a bicycle, or go there on foot”</td>
<td>S1 said that there is a relationship between the age and the transport means that people can use. In this way, the age indirectly influences the distance that people can travel. Although this could be true, it cannot support the hypothesis proposed in line 2. It seems as if S1 did not understand what TB1 aimed for.</td>
</tr>
<tr>
<td>5</td>
<td>TB1</td>
<td>“Nearer… uhmmm… I think you should know where are other bakeries and supermarkets which offer bread in the surroundings of Groesbeek and Beuningen”</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>S2</td>
<td>“Yes, but we investigated three Bakker Bart Bakeries, and not all bakeries”</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>TB1</td>
<td>“But maybe there are just not so many bakeries around Groesbeek. You could check that. How many bakeries? And where are they?”</td>
<td>TB1 told S what he was after at. However, he did not relate the distribution of bakeries to the size of market areas of those bakeries.</td>
</tr>
</tbody>
</table>

Notes: TB1 = Teacher B1; S1 = Student 1; S2 = Student 2.

11.2.4.6 Evaluation of the evaluation phase

Only one evaluative whole-class discussion was held: in Test II. In this discussion, the researcher tried to list the phenomena that influence the size of market areas of services in general, thereby asking the students for input. Students were asked to propose phenomena and explain the mechanism of the influence (Table 11-11). The discussion lasted only a couple of minutes. Table 11-12 presents a transcript from the whole-class discussion. Immediately after the evaluation phase, the researcher and teacher argued that they had the feeling that the whole-class discussion was not very effective, although they could not exactly say why at that time. Later on in the design process, it became clear that the initial theory was not legitimate, and that this was the reason why it was difficult to organize an effective whole-class discussion.
Chapter 11: Thick description

Table 11-11: Screenshot of the blackboard at the end of the whole-class discussion in Test II²¹

<table>
<thead>
<tr>
<th>Services:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Assortment</td>
</tr>
<tr>
<td>2. Price/quality ratio</td>
</tr>
<tr>
<td>3. Accessibility (parking facilities)</td>
</tr>
<tr>
<td>4. Atmosphere (friendly staff)</td>
</tr>
<tr>
<td>5. Next to other services</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customers:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Whether the visit is planned or by accident</td>
</tr>
<tr>
<td>2. Transport means</td>
</tr>
<tr>
<td>3. The maximum distance people are willing to travel</td>
</tr>
<tr>
<td>4. Income</td>
</tr>
</tbody>
</table>

Table 11-12: A transcript from the whole-class discussion in Test II²¹

<table>
<thead>
<tr>
<th>Line</th>
<th>Actor</th>
<th>Text / Action</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R</td>
<td>“We now have four phenomena on the side of the services [...] Are there any other phenomena on the side of the services?”</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>S1</td>
<td>“The location”</td>
<td>S1 mentioned the phenomenon ‘location’. It should actually be ‘situation’. This is a phenomenon at a high abstraction level.</td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>“Well, uhm... Could you explain that?”</td>
<td>R tried to specify the relevant phenomena (to change to a lower abstraction level).</td>
</tr>
<tr>
<td>4</td>
<td>S1</td>
<td>“Whether it is located in the centre of Nijmegen or, well, somewhere else”</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>R</td>
<td>“And how does that work?”</td>
<td>R tried to stimulate S1 to make the relevant relationship (at a lower abstraction level) explicit.</td>
</tr>
<tr>
<td>6</td>
<td>S1</td>
<td>“If you are in the Centre, the accessibility is not very good. But then there will be lots of people who visit your store because they are passing by”</td>
<td>S2 implicitly mentioned that people prefer to combine visits in order to save travel time and increase their chances for success. This is the causal mechanism (driving force) behind the relationship between size of market areas (of services) and the number of complementary competitive services in the vicinity (of those services).</td>
</tr>
<tr>
<td>7</td>
<td>S2</td>
<td>“They go to different shops”</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>R</td>
<td>“Do you mean that people combine visits to different stores? Yes. That is a good remark. And there is also... Well, some customers visit the store because they are just passing by. While other services are mainly visited by customers who come there to buy something specific. More like a planned trip”.</td>
<td>R did not recognize the relationship in students’ reasoning, and did not stimulate students to structure and expand their geographic thinking.</td>
</tr>
<tr>
<td>9</td>
<td>R</td>
<td>writes on the blackboard ‘whether the visit is planned or by accident’</td>
<td>R wrote down an ill-defined phenomenon that has no direct relationship with the size of market areas of services.</td>
</tr>
</tbody>
</table>

Notes:  R = Researcher; S1 = Student 1; S2 = Student 2.
11.2.4.7 Evaluation of the domain-specific construct for use in educational settings

In the evaluative talks, the teachers and researcher argued that the requirement that students conduct and inquiry project on the topic of services and customers did not provide enough guidance to ensure that students formulate good geographic inquiry plans. Students clearly need more guidance. Also, the theoretical insights from the geographic literature were not adequate for raising students’ geographic thinking to a higher level. The geographic theories seemed to have only weak links with the relationships in students’ geodata. Besides this, they were difficult to operationalize, for both students and for teachers. It was therefore concluded that there was a need for a good domain-specific construct for use in educational settings. Such a construct should consist of theoretical part (knowledge in the form of a theory for use in educational settings) and methodological part (knowledge about inquiry strategies for use in education settings). The construct could be used to design tasks, and to diagnose problems and shape support interventions when these tasks are conducted in the classroom. The question is: What are the characteristics of a good domain-specific construct for use in educational settings for the project ‘Services and Customers’? The following cycles of the design process will focus on, among other things, this question.

11.3 Cycle III

11.3.1 Introduction III

This section discusses Cycle III of the design process (Figure 11-7), in which the design of the inquiry project was revised, tested, and evaluated again. The main focus of Cycle III was on the design of the domain-specific construct for use in educational settings, the design of the preparation phase and the inquiry-planning phase, the design of the GIS phase, and the design of the presentation phase.

Teachers D1, D2, and D3 joined the research team. All decisions regarding the design process were made by the teachers and the researcher. The researcher revised the handouts. The tests were conducted in three classes at two schools (see Table 11-2): a 5VWO class at school A (Test III\(^A\)) and a 4HAVO class (Test III\(^D_{1+2}\)) and a 5VWO class at school D (Test III\(^D_3\)). In each of the three tests, the teacher conducted the inquiry project, with help from the researcher. Teacher A conducted the project for the second time, while teachers D1, D2, and D3 conducted the project for the first time. The data was analysed and interpreted by the researcher, and the results and interpretations were written down in a Research Report 2.
11.3.2 Design Stage III

11.3.2.1 Revision of the domain-specific construct for use in educational settings

In order to construct a theory about the size of market areas, teacher B1 and the researcher analysed the presentations of Test II\(^{B1}\) and tried to identify the factors that can explain the differences in the size of the market areas. This resulted in a theory in the form of a list of phenomena (Table 11-13), hereafter called ‘the preliminary theory’. At the time of Cycle III, the teacher and researcher thought that these phenomena could be seen as ‘characteristics of the service’ or ‘characteristics of the customer’. They were not certain whether the preliminary theory was consistent: they had the feeling that it did not include all relevant phenomena; and that it did not have a logical structure. Still, the research team decided to use the preliminary theory in Test Stage III.
The next question is: Which inquiry strategies are likely to result in inquiry with a high domain-specific quality? On the basis of the experiences of Test Stage II, the research team identified two standard inquiry strategies, which are hereby called ‘describing and interpreting the relationship between the size of market areas of services and another property of services’ and ‘comparing and contrasting the size of market areas of services and explaining the differences’. These standard inquiry strategies are specifications of the more general inquiry strategies ‘describing and interpreting the relationship between two properties of the same class’ and ‘comparing the properties of entities and explaining the differences’. The first connects to a topical-geographic approach; and the second to a regional-geographic approach (although the operations are performed on geo-referenced objects instead of regions). Descriptive and explanatory inquiry questions are part of both standard inquiry strategies.

Examples of good inquiry questions for the first standard inquiry strategy are: “What is the relationship between the assortment (of supermarkets) and the size of market areas (of those supermarkets)?”; and “Why is there a relationship between the two properties?” These questions are examples of, respectively, primitive key question 1-5-1 and 3-5 (see Section 5.7). Appropriate hypotheses are: “The assortment (of supermarkets) influences the size of market areas (of those supermarkets)”; and “There is a relationship between the assortment (of supermarkets) and the size of market areas (of those supermarkets) because people generally prefer to combine their shopping needs”. In order to answer the questions and test the hypotheses, students should select services of the same category that are similar for every property except one, in this case the assortment. Students could test the second hypothesis via survey questions such as: “Is your choice for visiting this supermarket determined by the assortment?”; and “Do you combine the visit to this supermarket with a visit to other shops?”

Examples of good inquiry questions for the second standard inquiry strategy are: “What are the differences in the size of the market areas of the Aldi supermarket, Albert Heijn supermarket, and Rio de Bio supermarket in Nijmegen?”; and “Which factors can explain the differences in the size of the market areas?”, which are examples of, respectively, primitive key question 2-1 and 1-1-2 (see Section 5.7). Appropriate hypotheses are: “The Rio de Bio has a large market area, as it is a very specialized supermarket with a small clientele and there are few similar supermarkets in the surroundings”; “The Aldi has an intermediatedly large market area, as it is a
cheap supermarket, and people with a low income are generally willing to travel further to go to a cheap supermarket”; and “The Albert Heijn has small market area, as it is an expensive supermarket, and as it is situated in a neighbourhood with a high per capita income, so it will draw customers from that neighbourhood.”

In the first standard inquiry strategy, it is more difficult to make a good selection of services than in the second standard inquiry strategy. However, in the second standard inquiry approach, it is more difficult to formulate good hypotheses and construct valid survey questions to test these hypotheses.

The inquiry plans from Test Stage II were analysed in order to get an idea of which standard inquiry strategy is best for the inquiry project. Teachers A, B, and C selected the inquiry plans which, according to them, had a high domain-specific quality. Eight out of, in total, 27 inquiry plans were selected. Analysis of these eight inquiry plans showed that all of them contained inquiry questions, hypotheses, and survey questions which fitted more or less into the standard inquiry strategy ‘comparing and contrasting the size of market areas of services and explaining the differences’. A couple of these inquiry plans also contained aspects of the other standard inquiry strategy. For example, one student pair said that: “By investigating three telephone shops in one street, we want to shut out the effect of location.” Although teachers A and C may have directed students implicitly into following the standard inquiry strategy ‘comparing and contrasting the size of market areas of services and explaining the differences’ in the introduction phase, the results of the analysis do indicate that students probably find it easier and more motivating to follow this standard inquiry strategy, and that it connects better to their way of thinking.

On the initiative of the researcher, it was decided to specify the learning goals regarding the development of geographic inquiry knowledge. Students should develop knowledge about the characteristics of inquiry strategies that are likely to result in inquiry with a high domain-specific quality, and learn to critically evaluate the domain-specific quality of their inquiry strategy. From now on, the term ‘deep learning about geographic inquiry methods’ is used for this learning goal.

### 11.3.2.2 Revision of the design of the introduction phase

In order to avoid problems caused by the learning difficulty that students often do not have a clear idea about what they are supposed to do and how they are supposed to do it, the research team conjectured that it is effective to show them the entire process from collecting geodata to constructing maps in a PowerPoint presentation than to show only the end products, as was done in Test II. As the teachers were afraid that the students would copy the survey questions in the presentation, they decided to use an Excel file with geodata about the students of their school. This file consisted of two sheets, one with the geodata per student (name, postcodes, class, gender, and profile) and one with the aggregated geodata for the school. The PowerPoint presentation contained sheets which showed how to convert the postcodes of students to X and Y coordinates, calculate the absolute distance for every student, aggregate the geodata, add the geodata tables in GIS, convert the geodata tables to map layers, and change the symbology of the map layers. So the PowerPoint presentation would show students how to construct maps that are similar to the maps that they are supposed to construct, on the basis of geodata that is similar to the geodata they are supposed to collect.
11.3.2.3 Revision of the design of the preparation phase

On the basis of the experiences in Test Stage II, the research team argued that students need to have more training in order to be able to formulate good surveys. The questions are: Which types of survey questions are useful in this project? And what are the characteristics of correct survey questions? Together, teacher B1 and the researcher constructed a list of eight types of survey questions (Table 11-14) and a list of seven requirements for good survey questions (Table 11-15). It was conjectured that students construct inquiry plans with a higher domain-specific quality if they learn about the different types of survey questions and the requirements of correct survey questions beforehand. A preparatory assignment was therefore included in the preparation phase. This lesson consisted of two assignments. The plan was that in the first assignment, students would have to categorize a given set of survey questions with the help of the list of eight types of survey questions. In the second assignment, students would have to identify and classify errors in another set of survey questions with the help of the list with the seven requirements of correct survey questions. It was conjectured that this preparatory assignment is an effective way to avoid the problems that students’ inquiry plans often contain the wrong type of survey questions and incorrect survey questions.

Table 11-14: The eight types of survey questions

<table>
<thead>
<tr>
<th>No.</th>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Factual multiple choice questions (one answer possible)</td>
<td>“What is your travel time from home to work (0-15min / 15-30min / &gt;30min)?”</td>
</tr>
<tr>
<td>2</td>
<td>Factual multiple choice questions (multiple answers possible)</td>
<td>“Which parties have you ever voted (the Green Party / the Liberal Party / the Social Democrats / the Christian Democrats)?”</td>
</tr>
<tr>
<td>3</td>
<td>Evaluative multiple choice questions (feedback on a proposition)</td>
<td>“What do you think about the proposition ‘the migration policies should be more strict’ (fully disagree / disagree / disagree nor agree / agree / fully agree)?”</td>
</tr>
<tr>
<td>4</td>
<td>Evaluative multiple choice questions (opinion scale)</td>
<td>“What do you think about the migration policies (too strict / a bit too strict / just right / a bit too loose / too loose)?”</td>
</tr>
<tr>
<td>5</td>
<td>Evaluative multiple choice questions (opinion order)</td>
<td>“How do you value the assortment of this supermarket on a scale from 0 to 10?”</td>
</tr>
<tr>
<td>6</td>
<td>Evaluative hierarchic questions</td>
<td>“What are the three most important annoyances about this neighbourhood (noise on the streets / dirty streets / bad houses / lack of shops / lack of parking space)?”</td>
</tr>
<tr>
<td>7</td>
<td>Open questions (number)</td>
<td>“How often do you visit the city centre?”</td>
</tr>
<tr>
<td>8</td>
<td>Open questions (description)</td>
<td>“How do you feel about the plans for renovation of the city centre?”</td>
</tr>
</tbody>
</table>
11.3.2.4 Revision of the design of the inquiry-planning phase

The research team decided to stimulate students implicitly in one-on-one (student and teacher) discussions to follow the standard inquiry strategy ‘comparing and contrasting the size of market areas of services and explaining the differences’. This probably helps them to construct clearer inquiry plans with good hypotheses and a strong spatial perspective, and in such a way overcome β problems caused by the six inquiry planning-related learning difficulties mentioned in Section 11.2.4.2. Teachers D1 and D2, who would conduct the inquiry project with a 4HAVO class, even decided to instruct all students explicitly to follow this standard inquiry strategy. They would give students the standard inquiry questions and hypotheses, and instruct them to customize the standard inquiry questions and hypotheses.

11.3.2.5 Revision of the design of the Excel phase

It was conjectured that the number of α problems as a result of Excel files that contain geodata tables that do not meet the criteria of a proper database table\(^{ap9}\), and Excel files that are incomplete\(^{ap9}\), could be reduced in three ways. First, it could be reduced by telling students about the criteria of a proper database table, and by explaining to them why it is important to construct geodata tables that meet these criteria\(^{NC28}\). Second, it could be reduced by instructing students clearly to read the worked-out examples in the Excel handout very precisely and to follow it right to the end\(^{NC29}\). Third and finally, it could be reduced by including a task at the end of the Excel handout in which students have to check their Excel files\(^{NC30}\). As spaces are often difficult to identify, students should use the find and replace tool (CTRL + H) to replace all spaces by underscores.

11.3.2.6 Revision of the design of the GIS phase

The research team decided to revise the structure of the GIS phase completely and follow Option 3 instead of Option 2 (see Table 11-4). In the revised design, students would follow a handout with step-by-step instructions with worked-out examples that show which operations to perform and how to perform these operations (compare, for example, Figure 11-5 with Figure 11-8). It was conjectured that such a structure is more viable to ensure that students are able to visualize their geodata without encountering α problems\(^{NC14}\).

---

Table 11-15: The seven requirements of correct survey questions

<table>
<thead>
<tr>
<th>No.</th>
<th>Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The questions should not be suggestive</td>
</tr>
<tr>
<td>2</td>
<td>The questions should not be too personal</td>
</tr>
<tr>
<td>3</td>
<td>The questions and answers should be clear and unambiguous</td>
</tr>
<tr>
<td>4</td>
<td>The number of answers should be limited</td>
</tr>
<tr>
<td>5</td>
<td>The multiple choice answers should have a logical order</td>
</tr>
<tr>
<td>6</td>
<td>The multiple choice answers should exclude each other</td>
</tr>
<tr>
<td>7</td>
<td>The multiple choice answers should cover the entire range of possible answers</td>
</tr>
</tbody>
</table>
Enquiry data with a number format (such as the average age of the customers) can be visualized in a quantities map with large symbols for high values and small symbols of low values. If your file contains this kind of data, go to Step 26. If your file does not contain this kind of data, go to Step 28.

**Figure 11-8: Fragment of the GIS handout used in Test Stage III**

Step 26: Copy one of the pie chart map layers (see the figure below).

- **Step 26a**
  Click (right button) on the name of one of the pie chart map layers and choose for ‘copy’.

- **Step 26b**
  Click (right button) on ‘layers’ and choose for ‘paste’.

Step 27: Change the symbology of the new (top most) pie chart map layer and construct a quantities map layer (see the figure below).

- **Step 27a**
  Click (right button) on the name of one the new (top most) pie chart map layer and choose for ‘properties’

- **Step 27b**
  Click on the symbology tab.

- **Step 27c**
  Choose ‘quantities maps’ and ‘graduated symbols’

- **Step 27d**
  Choose the right attribute (column header in the table): the column with average values.

- **Step 27e**
  Check this box.

- **Step 27f**
  Click on the OK button.
11.3.2.7 Revision of the design of the presentation phase

In order to overcome the β problem that students include few notes of reflection on the domain-specific quality of their inquiry project in their presentations and reports, it was decided to replace the note: “What went well and what did not go well?” in the requirements of the presentations and report by the note: “Reflect on your inquiry strategy. Did you select the right services, and did you formulate the right hypotheses and survey questions? What would you do differently if you had to do the same project all over again?” The research team conjectured that it helps to tell students how to reflect in order to stimulate them to reflect more deeply on the domain-specific quality of their inquiry project.

11.3.2.8 Revision of the design of the evaluation phase

It was decided to split the class in two small groups, so that there would be enough time for an evaluative whole-class discussion in which the teacher and students summarize and abstract the knowledge that was gained. In the discussion, the teacher would list the phenomena that influence the size of market areas of services in general, thereby asking the students for input. The teachers would use the preliminary theory about the size of market areas of services (see Table 11-13) to structure students’ input. Teachers D1 and D2 decided not to let students present their work and not to organize an evaluation, as their schedule did not allow time for such activities.

11.3.3 Test Stage III

11.3.3.1 Realized teaching and learning

The introduction phase, preparation phase, research planning phase, and data collection phase went as planned. However, the Excel phase did not go as planned. In Test III, the teacher had told the students to anticipate the next lesson and enter their geodata in Excel at home. In the lesson, the teacher handed out the Excel handouts and instructed the students to follow the handout and change the structure of the geodata tables so it would meet the requirements of a proper database table. This did not work. Students often skipped parts of the Excel handout, which resulted in many α problems in the GIS phase. On the basis of this experience, it was concluded that students should do the geodata entering, manipulating, and aggregating at school, under the supervision of a teacher.

The GIS phase did go as planned, although many α problems occurred, especially in Test III. The pressure on the teacher (and researcher) was somewhat lower than in Test Stage II. However, it was still high. Despite these problems, the teachers were enthusiastic about the output of the inquiry project. The inquiry project about the gyms in Gorinchem (see Section 10.2) is an example of one of the inquiry projects that were carried out in Test Stage III.

11.3.3.2 Collected data

Four types of data were collected during and shortly after the tests. First, the inquiry-planning phase of Test III and II was videotaped, with a focus on instances in which students were formulating inquiry plans together and instances in which the students and teacher were discussing the inquiry plans. Second, field notes were made during the GIS phase. All α
problems were registered. Third, the inquiry plans, presentations and reports constructed by students were collected. Fourth and finally, surveys were conducted with the students immediately after they received their grades.

11.3.4 Evaluation Stage III

11.3.4.1 Evaluation of the inquiry-planning phase

Fewer students asked for help in the inquiry-planning phase of Test Stage III than in the inquiry-planning phase of Test Stage II. On the basis of this experience, the research team concluded that most students understood what they were supposed to do and how they were supposed to do this. It was concluded that showing students the PowerPoint presentation with the worked-out examples was an effective way to avoid β problems caused by the learning difficulty that students often find it difficult to understand what they are expected to do, and how they can do it\textsuperscript{[110]}\textsuperscript{[21]}. However, there were still some students who asked questions about how they could show survey geodata in maps. In summary, the revisions in the design helped to avoid these β problems, but more is needed to prepare the students for their inquiry project\textsuperscript{[PAC15]}. The sheet with the requirements for the inquiry plans helped to improve the clarity and completeness of the inquiry plans, and to overcome the β problem that students often construct unclear and incomplete inquiry plans\textsuperscript{[BP5]}. Almost all inquiry plans followed the set-up in the sheet, and every inquiry plan contained a list with survey questions, including multiple-choice answers. The teachers in Test Stage III were therefore able to give more and better feedback on the domain-specific quality of the inquiry plans than the teachers in Test Stage II.

It seemed as if students used more different types of survey questions and formulated less questions that did not meet the requirements of correct survey questions than in Test Stage II. However, it is not certain whether this was because the students followed the preparatory assignment ‘constructing surveys’ beforehand, or because the teachers offered more support when students were constructing inquiry plans. This issue was therefore not investigated in detail, and it could not be tested whether students actually construct better surveys if they learn about the characteristics of the different types of survey questions and the requirements of correct survey questions beforehand\textsuperscript{[PC26] and C27].

The inquiry questions and hypotheses in students’ inquiry plans in Test Stage III were more spatial than those in Test Stage II. This indicates that it helps to stimulate students explicitly or implicitly to follow the standard research strategy ‘comparing the market areas of services and explaining the differences’. However, students still often formulated survey questions that were potentially ineffective for testing their hypotheses and still included few spatial key concepts in their survey questions. For example, the student pair who investigated the four gyms in Gorinchem asked the visitors of the gyms for their age and visiting frequency. However, they did not link the differences in these properties to the differences in the size of the market areas. They just included the two survey questions because they hypothesized that there would be differences in the visiting frequency and average age. So the external validity and the spatial character of the inquiry plans were still not very high. The videodata show that one-on-one discussions between students and teacher rarely focused on this issue. In the evaluative talks, the teachers said that they found it difficult to support students in constructing good surveys, because they had only intuitive ideas about which questions are suitable.